**Tableau Prep 簡易使用手冊 -- 使用進線數資料為清理範例**

1. 開啟檔案

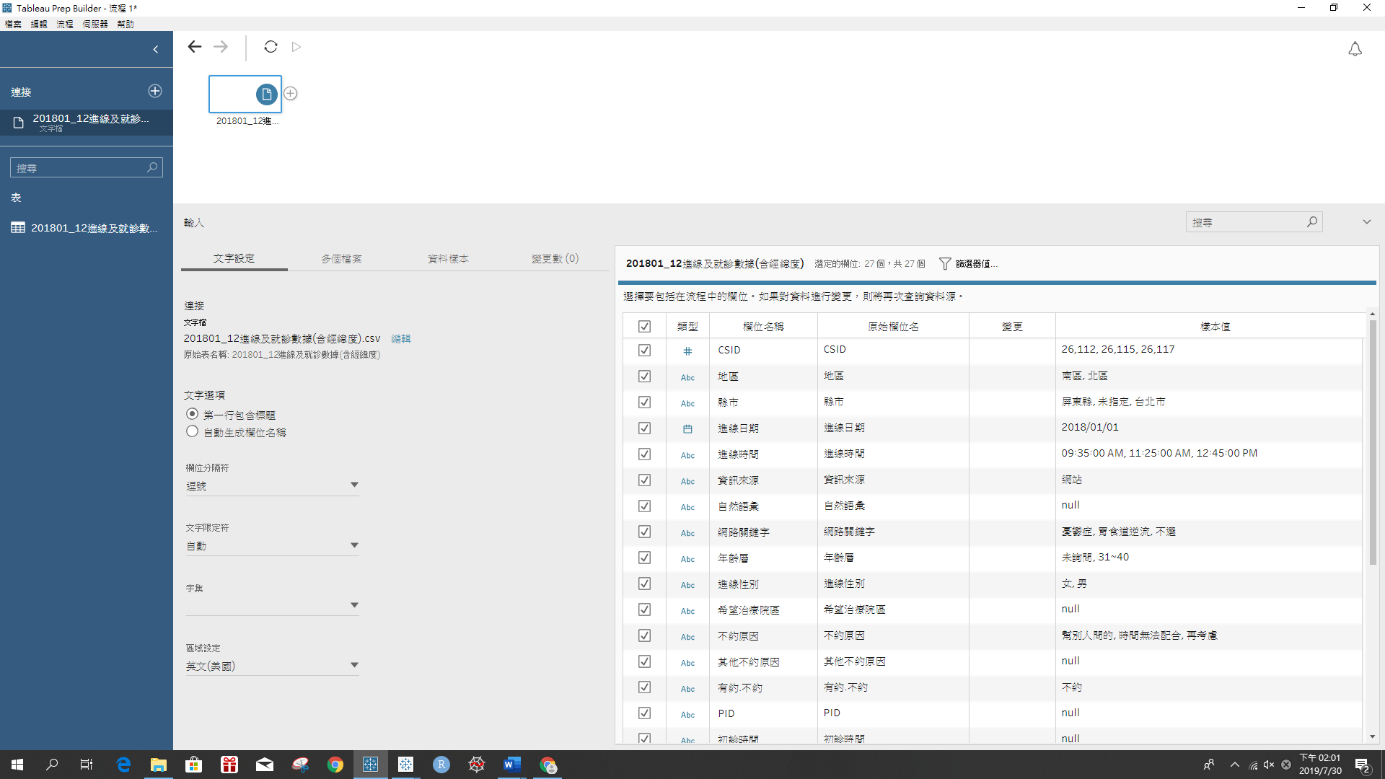
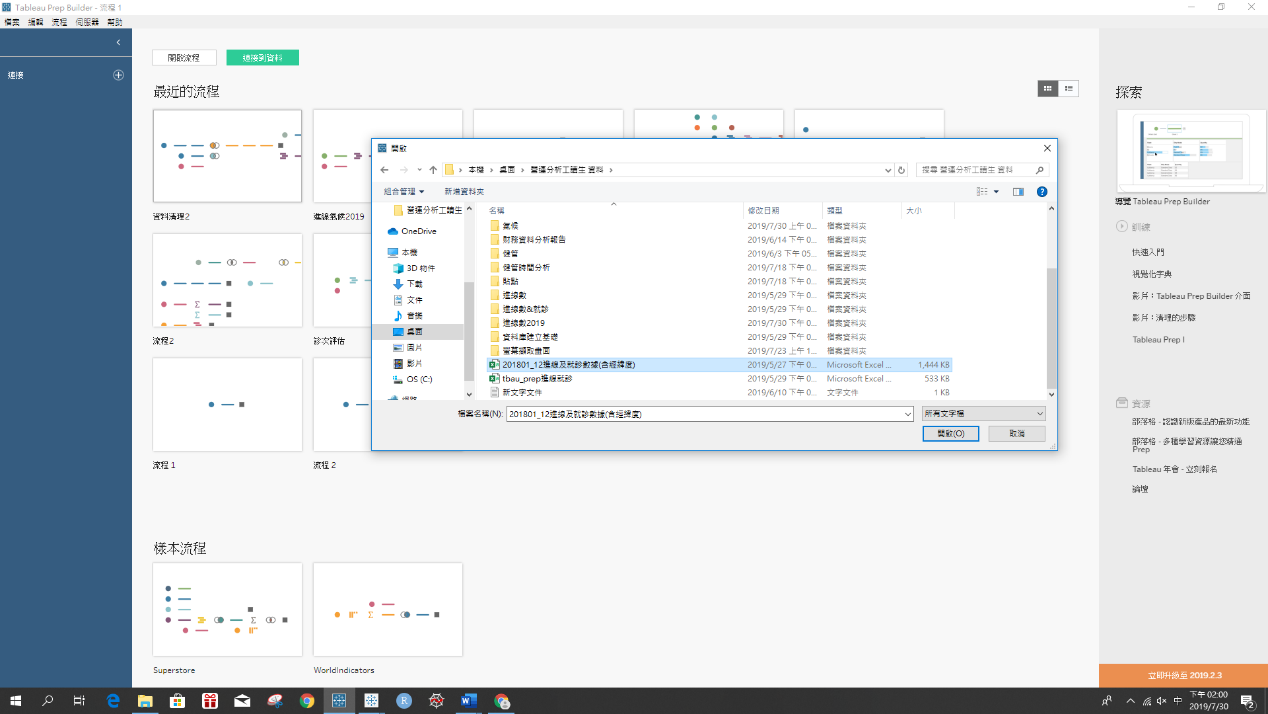
一張含有 螢幕擷取畫面 的圖片

自動產生的描述

一張含有 螢幕擷取畫面 的圖片

自動產生的描述

Table prep 可以串接不同來源與格式資料，其中最常用的是Excel格式與csv(文字檔)格式；另外，也可以連接伺服器抓資料清理。



資料表

資料源

工作流程圖

工作區

我們選擇文字檔，並且匯入進線資料，匯入後我們可以看到資料被預先丟入工作區。

Tableau Prep 操作介面可以分為四大區塊 :

* 資料源

代表匯入資料來源，我們再清理一份分析資料，可能會需要串接不同的資料來源，來擷取我們想要分析的資料內容。若需要新增資料源，只需選擇⊕匯入資料即可。

* 資料表

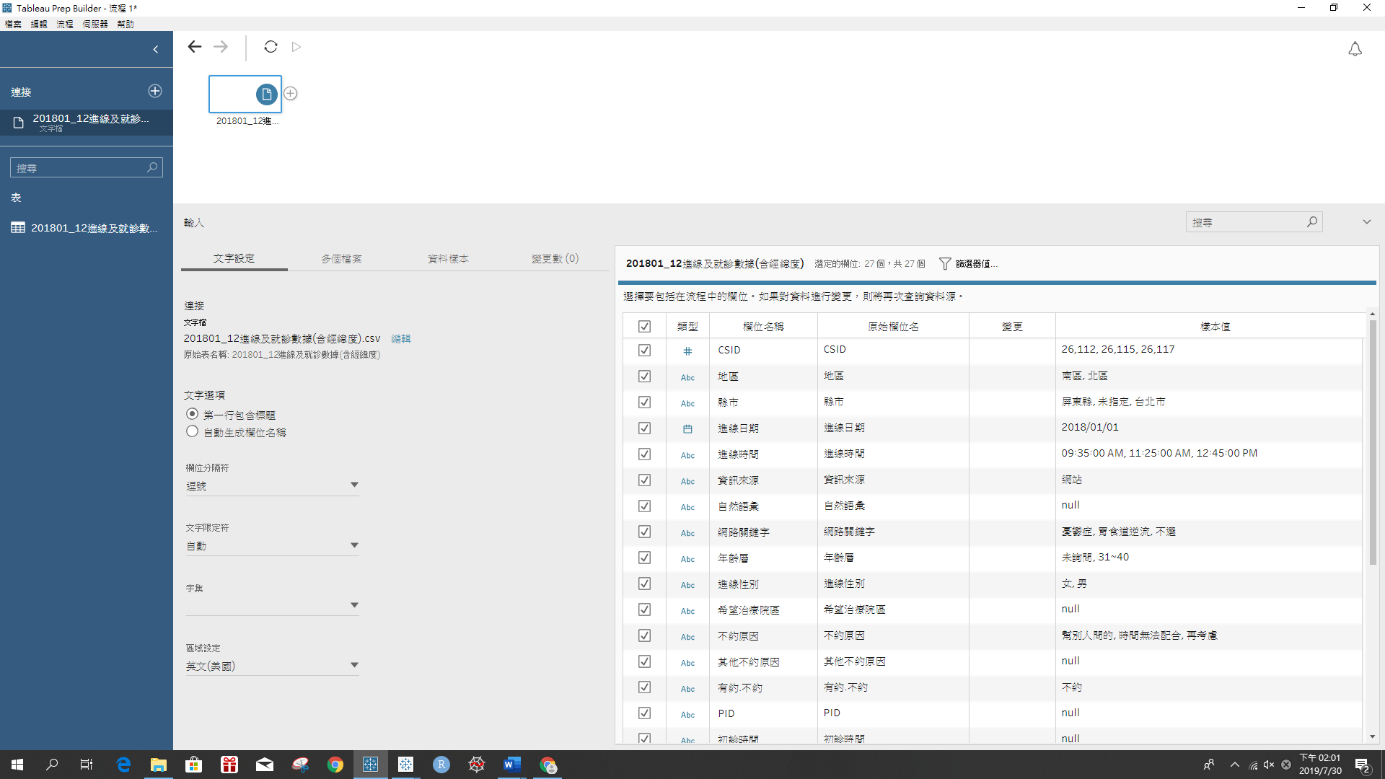
代表該資料源目錄下的資料表，若以Excel格式來說，也代表一個檔案中的活頁夾資料表。若要清理特定資料表，將之點選拖曳到工作流程圖區域即可。

* 工作流程圖

任何對資料的步驟都會依「視覺化」的流程圖紀錄於檔案中，工作流程圖是步驟的模組，今天若以相同格式之資料表取代原始資料表，透過相同的流程也可以跑出相應的結果。

工作流程圖右上方

* + 左箭頭 : 上一上
  + 右箭頭 : 下一步
  + 雙箭轉換 : 更新原始資料表內容 (若資料表原數據有更正)
  + 三角形 : 運行流程 (用於產出清理結果)



從以上工作區我們可以再細分為數個功能 :

* 文字設定

確定匯入之格式

* 多個檔案

選擇聯集方式

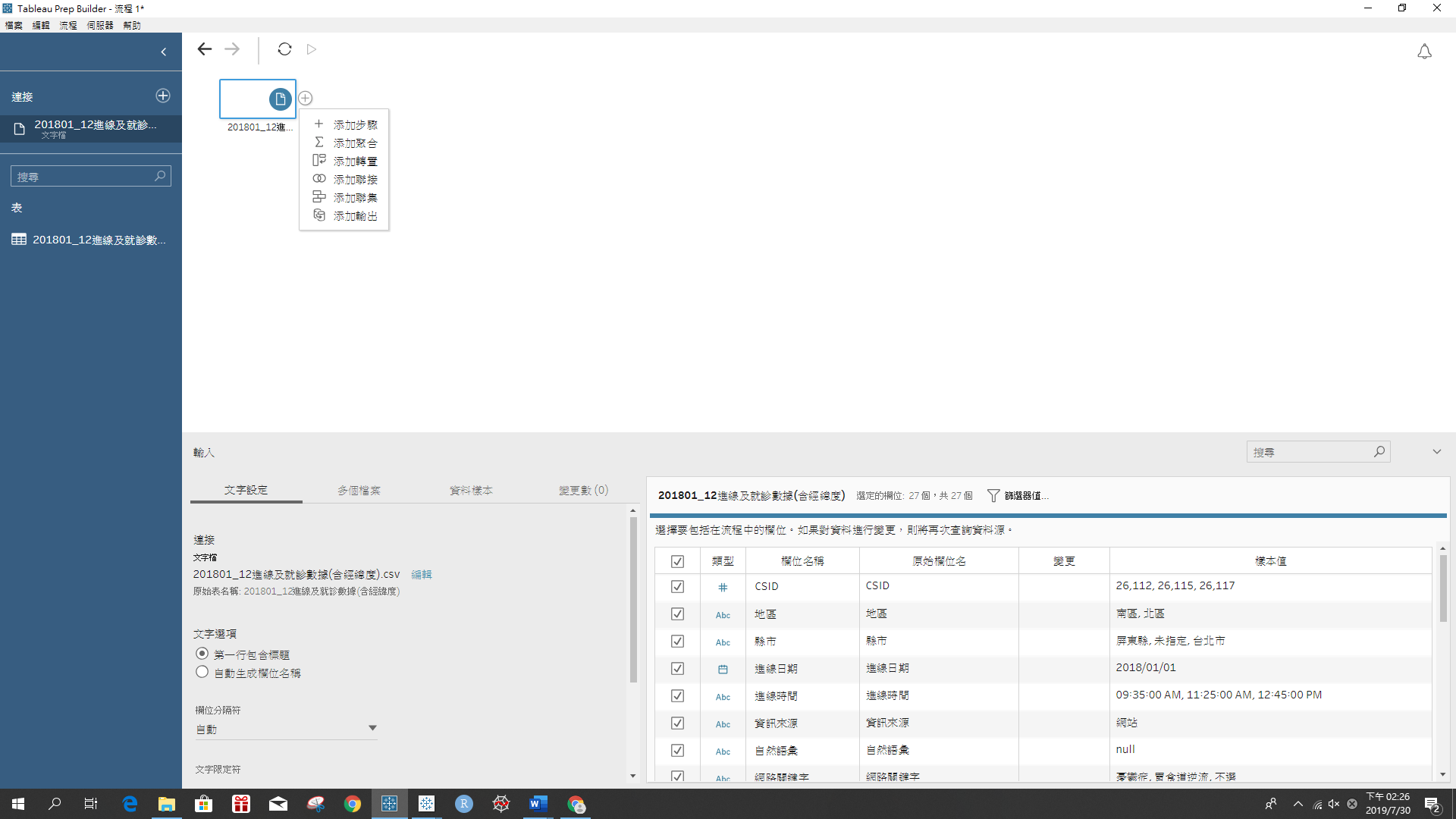
* 資料樣本

用於大資料或統計資料

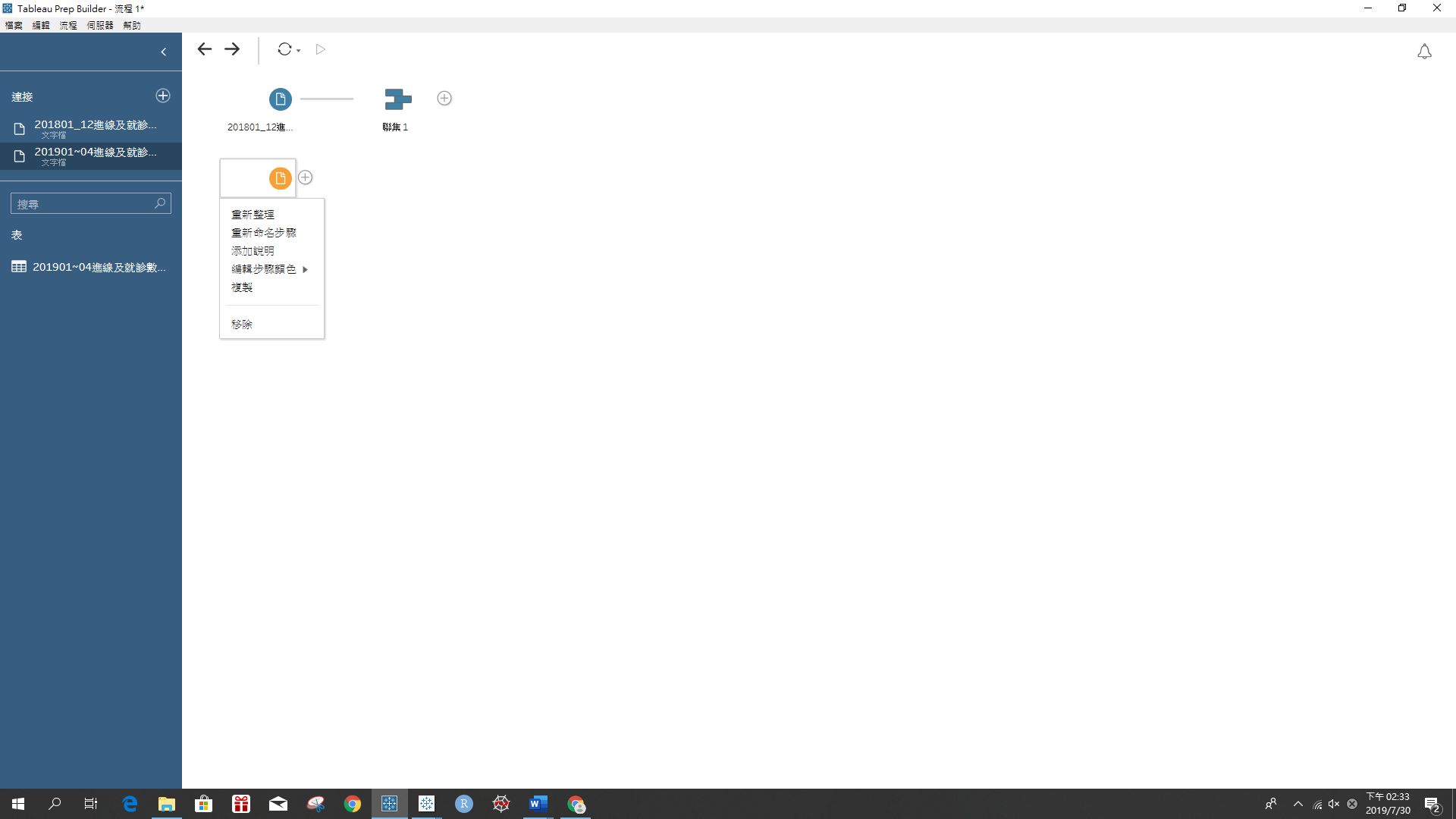
* 變更數
* 選擇欄位

勾選需要清理之欄位

1. 清理步驟
   1. 聯集



選擇清理流程檔案旁的⊕選擇要操作的步驟，若我們想要聯集2019年資料，可以選擇添加聯集，或是將另外資料拖曳到原始資料上，選擇聯集。



點選資料本身右鍵，可以選改變清理流程名稱格式、改變顏色、添加說明或移除該資料表。

以下聯集資料

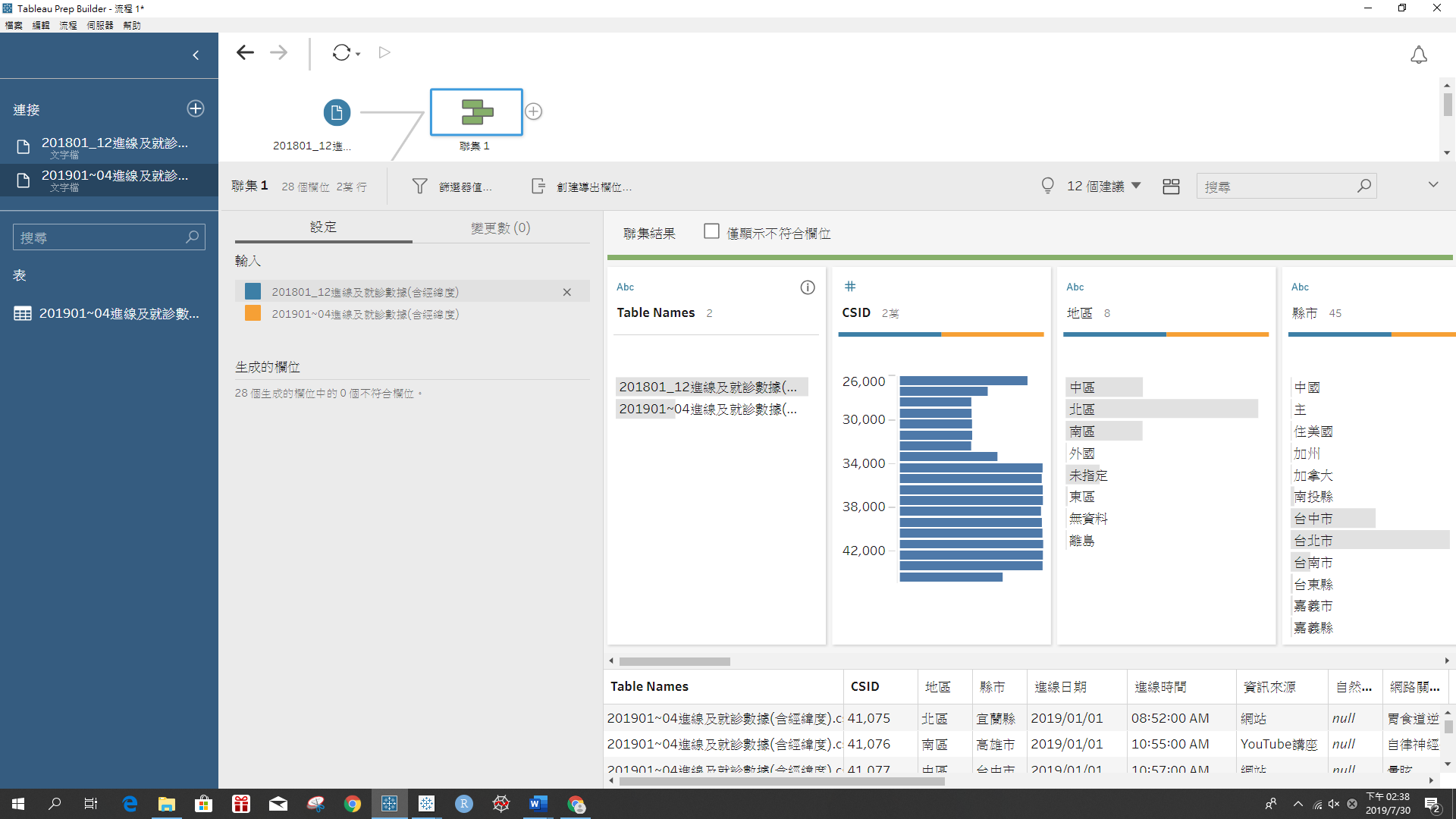
一張含有 螢幕擷取畫面 的圖片

自動產生的描述

一張含有 螢幕擷取畫面 的圖片

自動產生的描述

拖曳資料表，我們可以選擇聯接、聯集。聯接乃交集，是透過不同資料表相同的主鍵(ID)來添加資料，融合兩份資料、聯集是相同欄位格式之資料添加在一起。

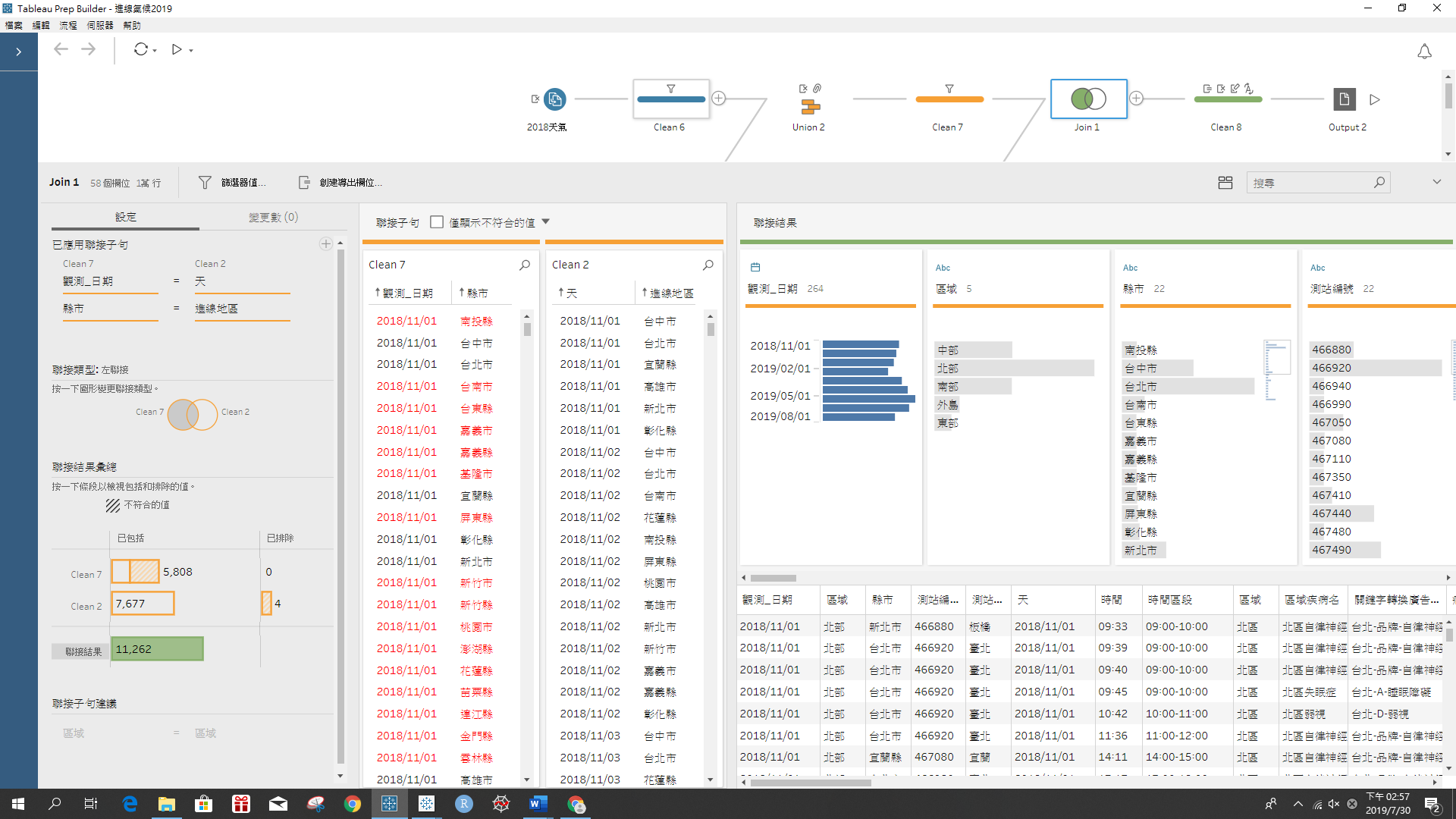
工作區左邊為聯集後資訊，右邊為資料狀況，每一個欄位資料卡上方色條代表源自於哪一個資料表，若該色條僅有一色代表其中一方資料表缺少該欄位，我們也可以透過選擇工作區上方僅顯示不符合欄位來檢視異常欄位。

資料卡下方為聯集後資料樣本。

* 1. 聯接

一張含有 電腦, 螢幕擷取畫面, 室內, 膝上型電腦 的圖片

自動產生的描述若要交集兩份資料，可以使用聯接，最快的方式將資料表拖曳至目標資料表上方選擇聯接。

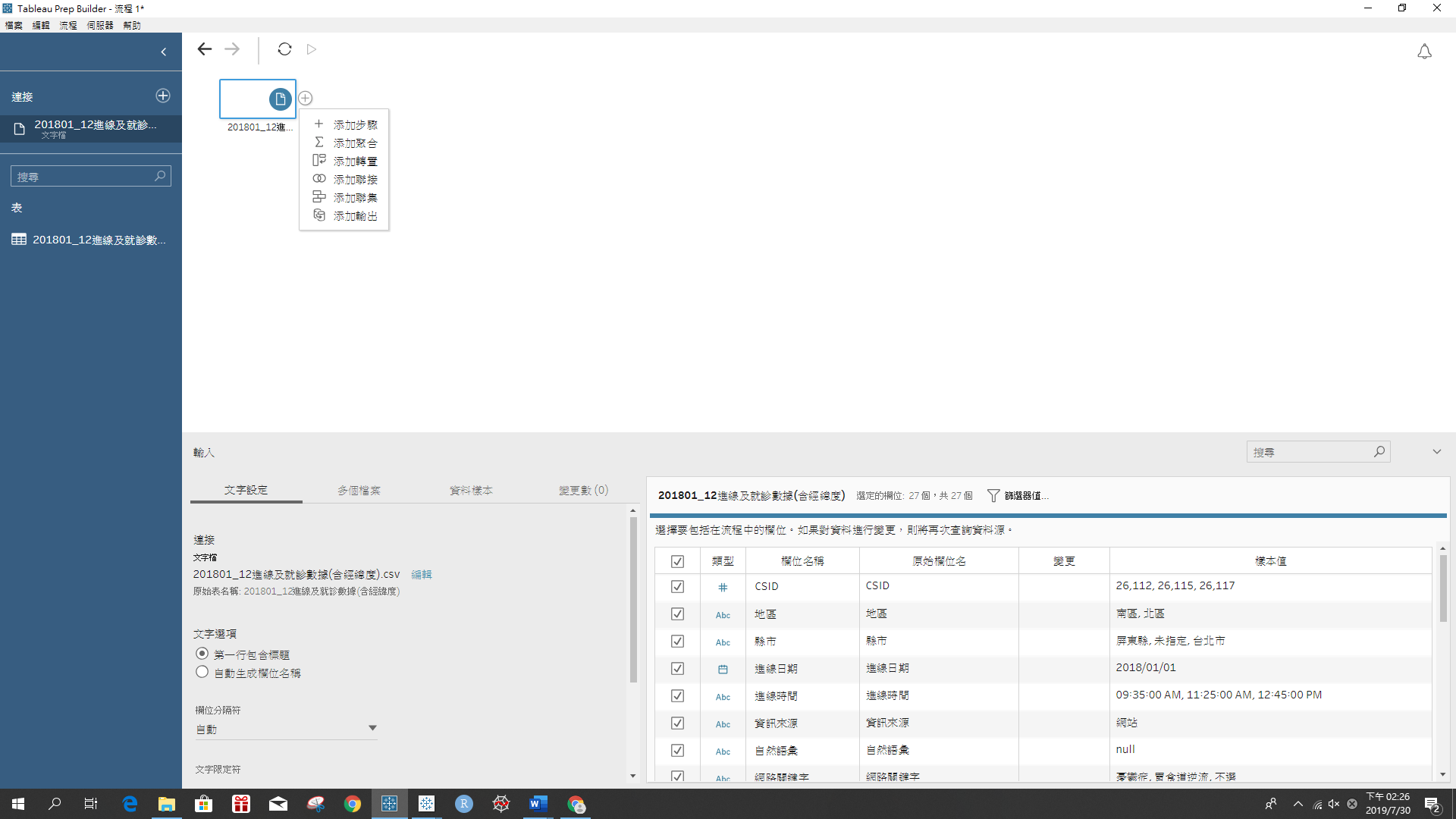


* 聯接子句 : 兩方資料主鍵ID，添加一個以上之子句讓聯接資料更精確。
* 聯接類型 : 選擇內接、外接、左接、右接，是資料應用情況而定。
* 一張含有 電腦, 螢幕擷取畫面, 室內, 膝上型電腦 的圖片

  自動產生的描述聯接會總結果 : 聯接已包括資料，深色斜線區為沒有配對之資料，若聯接則該配對欄位為NULL

紅字為不相配之資料，可以透過選擇僅顯示不符合的值來觀察資料是否有異。

* 1. 步驟



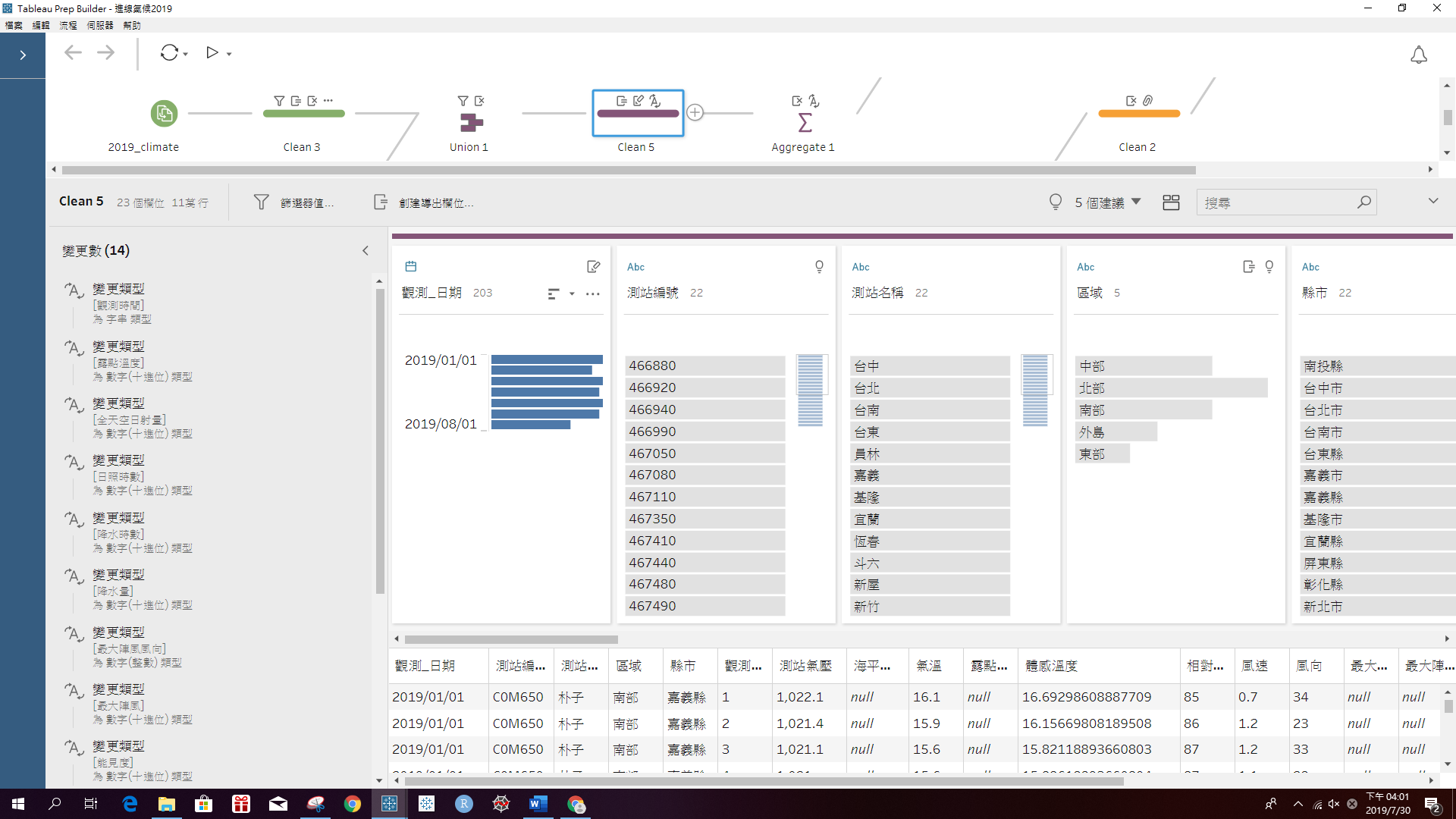
一張含有 螢幕擷取畫面, 電腦, 室內, 膝上型電腦 的圖片

自動產生的描述添加步驟，進入詳細資料清理

樣本資料

欄位資料卡

透過已有的資料計算一個新的欄位

一張含有 螢幕擷取畫面 的圖片

自動產生的描述

流程視覺化

步驟內流程，隨意點選某流程編輯，結果直接套用於所有流程。

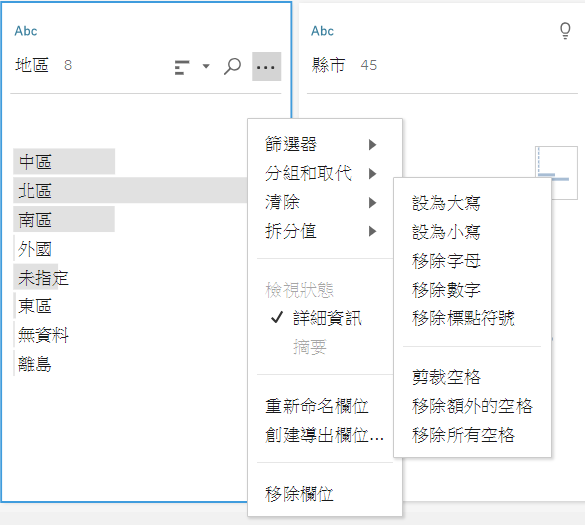
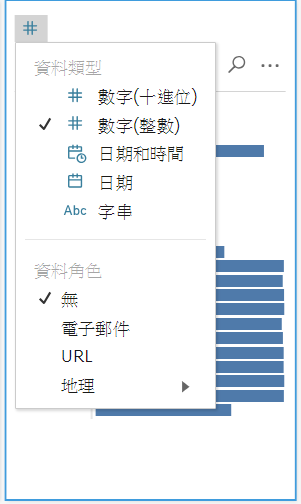
流程視覺化

* 欄位資料卡

編輯該欄位中的資料，包含轉換資料格式、清理遺漏值、清理異常值、群組資料、資料分割與取代等等…

一張含有 螢幕擷取畫面, 電腦, 室內, 膝上型電腦 的圖片

自動產生的描述



* 樣本資料

隨時反映清理過程中資料內容樣態

* 創建導出欄位

顧名思義，創立一個新的欄位，依目的添加不同的計算式導出不同結果。

一張含有 螢幕擷取畫面, 電腦, 膝上型電腦, 室內 的圖片

自動產生的描述一張含有 室內, 電腦, 螢幕擷取畫面, 膝上型電腦 的圖片

自動產生的描述一張含有 電腦, 室內, 膝上型電腦, 螢幕擷取畫面 的圖片

自動產生的描述

* 1. 聚合

於資料步驟右方⊕選擇添加聚合

一張含有 螢幕擷取畫面 的圖片

自動產生的描述

將要分組之類別資料拖入分組欄位，例如 : 縣市、性別

將要聚合之數值資料拖入聚合欄位，例如 : 人數、療程數

一張含有 螢幕擷取畫面, 電腦, 室內, 膝上型電腦 的圖片

自動產生的描述

於聚合欄位，每一個資料卡上方可以選擇聚合方式

一張含有 螢幕擷取畫面 的圖片

自動產生的描述

* 1. 轉置

於資料步驟右方⊕選擇添加轉置

一張含有 螢幕擷取畫面 的圖片

自動產生的描述將目標欄位拖曳進中間轉置的欄位

一張含有 螢幕擷取畫面, 電腦 的圖片

自動產生的描述

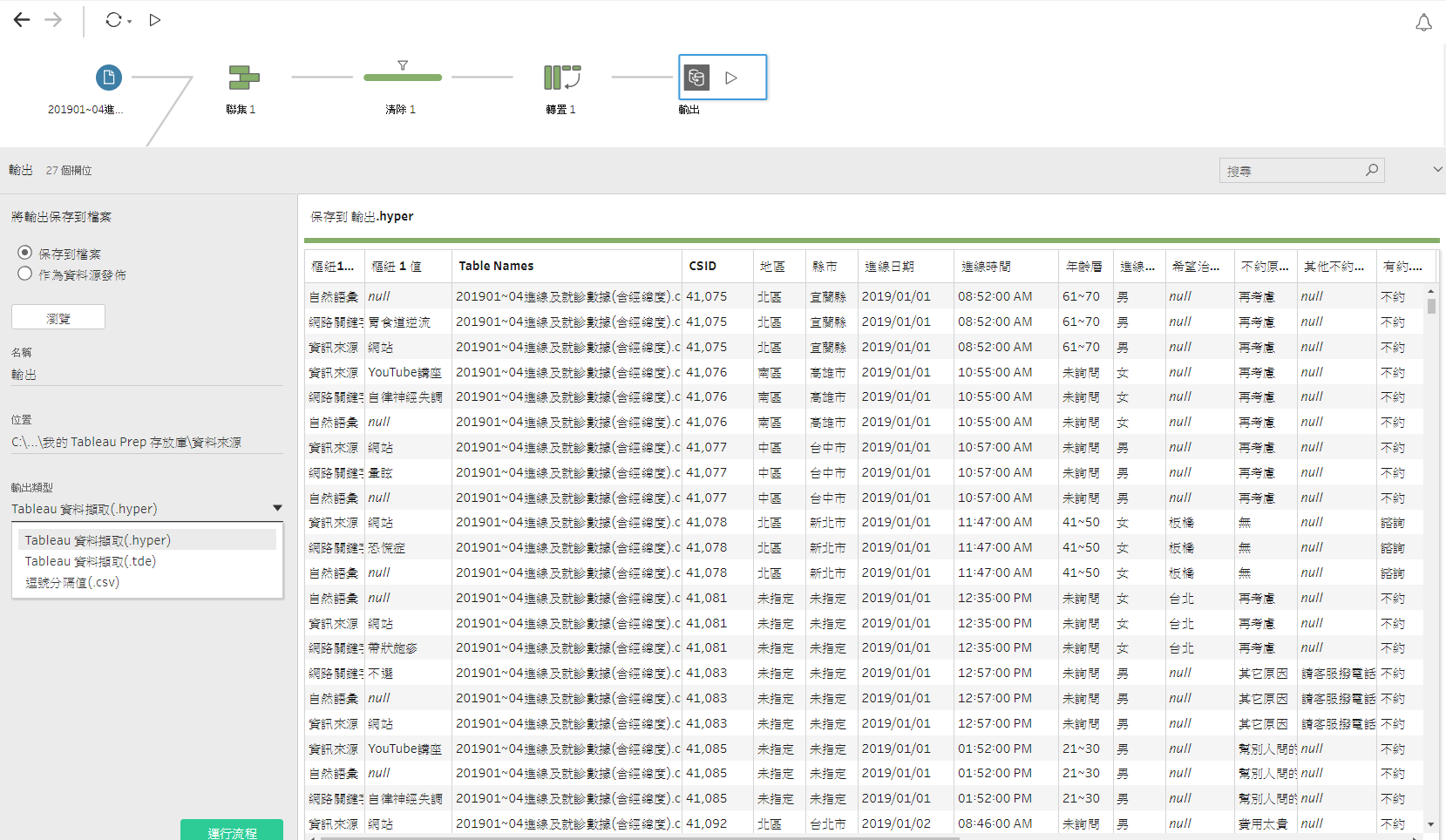
添加樞紐值

一張含有 螢幕擷取畫面, 電腦 的圖片

自動產生的描述

1. 資料輸出

於資料步驟右方⊕選擇添加輸出



* 將輸出保存到檔案，存檔；將輸出做為資料原發佈，發佈到伺服器。
* 瀏覽

選擇存檔位置和名稱

* 輸出類型
  + Tableau 資料擷取
    - .hyper
    - .tde
  + .csv (建議使用)
* 運行流程

存檔